Evans Mosomi

100719552

Lab - Data Storage Implementation: KV & Relational

VIDEO LINKS

Q:

- Sink connectors - This allows you to export data from Apache Kafka topics to any relational database via a driver.
- Source connectors - Ingests databases and stream table updates as a whole into Kafka topics.

  Advantages
  - Easy to configure - It takes advantage of the streamlined process of integrating Kafka with a system.
  - Scalability - Different connectors can be used to connect to as many external data sources
  - Troubleshooting is easier - error handling. Issues can be pinpointed to the specific connector where they failed and not the entire system as having a problem. This is because each topic works independently to the service it needs.

- **Availability** - Kafka connectors rely on the Connect framework that is Kafka-based. Microservices are created that are duplicates of the original service, and for each, a new connector is built and thus if one connector is down, other connectors are able to take over and keep the system active.

  Kafka converters
  - JSON - Creates several config options: object.additional.properties; use.optional.for.nonrequired; decimal.format
  - Avro - Requires key.converter and value.converter to serialize and deserialize data
  - Protobuf _ Has 3 components to package information: type, default, importance

Q:

- **Key-value database -** A non-relational database type using key-value pairs for data storage; whereby a key becomes the identifier and the value is the desired content.

- Advantages
  - Highly partitionable
  - Allows horizontal scaling
  - Uses compact index structures to locate values for O(n)
  - Real-time random access
  -Caching

- Disadvantages
  -No querying
  -Caching

- Examples
  Amazon DynamoDB
  MongoDB
  Redis

Q: Applications implemented using the sample dataset

- Object detection
- Geomapping
- Hydrodynamic modeling

VIDEO LINKS